



# **National Archives**

## **Digital Imaging Policy**

**February 2016**

## 1. Introduction

The National Archives digital imaging policy is designed to support and facilitate the preservation of archives in its care. A core statutory function of the National Archives is to preserve archives, and digital imaging is the primary means of assisting preservation whilst simultaneously enabling wider access to collections of national significance. This document details the National Archives policy on digital imaging with regard to the long term preservation of archives.

We will ensure that:

- All digital imaging is carried out in line with international best practice;
- There is a coherent approach to digitisation; and
- Access to digital collections will be provided.

[For the purposes of this document 'digital imaging', and 'digitisation' refers to both the physical process of scanning or photography, and the series of activities leading to the creation of the digital copy for use by the public.]

## 2. Background

The National Archives is entrusted with preserving and providing access to archival documents of the Irish State and its people.

This policy has been developed to:

- facilitate the increasing demand for access to archives from the public and to meet our obligations under the National Archives Act, 1986 in providing such access ;
- ensure that collections of high public interest are prioritised for digitisation;
- guarantee that standards and procedures are in place to ensure digitisation of archives is carried out in a systematic and uniform way with due regard for the care of the items;
- advise and inform the public and other parties as to our approach to digitisation.

## 3. Principles

We approach digital imaging with due regard to the selection, preparation and accessibility of the archival object.

Archive collections are selected for digitisation based on the following conditions:

- **Demand** - Proposals to digitise archive material should demonstrate that the documents are in high use or that digital availability would increase their use;
- **Significance** – Collections should have a national or research significance;
- **Scope** - An entire collection, or complete series from a collection should be identified rather than the selection of discrete items from various collections;
- **Condition** - The collection must be physically robust enough for digitisation. A conservation survey will be conducted and only records requiring low or medium levels of repair will be digitised;

- **Metadata** - The collection should be processed and listed in line with National Archives processing guidelines (listed in ISAD(G), IGAD or Dublin Core, and each item or page sub-numbered).

#### 4. Collaborative Projects

We will engage in collaborative projects with other organisations which will:

- Provide increased access to archival collections by the public;
- Promote the holdings of the National Archives;
- Leverage 'added value' by digitally linking collections;
- Develop competencies in digital humanities, open data, or similar initiatives.

#### 5. Conditions

The National Archives is statutorily bound to ensure the preservation and integrity of records in its care. Records will always remain in our custody and consequently we **do not** facilitate the following:

- Lending to a third party to facilitate digitisation;
- Digitisation of unprocessed collections. All collections must be formally accessioned, listed, sub-numbered and conserved (if necessary) prior to digitisation.
- Digitisation of collections where there are conservation concerns. A conservation survey will be carried out and consent from conservation staff obtained prior to digitisation.

#### 6. Reprographics or publication

As part of our services to the public we provide a reprographics service, providing hard copies, digital images and photographic prints on request.

- For large copy orders an assessment of the quantity of material to be copied will be carried out and if the collection is of high research interest, the entire series may be digitised and made available on the National Archives website.
- For one off images for publication purposes, a copy will be retained for one calendar month.
- If the material is deemed to be of high research value or likely to be the subject of similar requests, we may retain a copy for preservation purposes.

#### Legal

#### 7. Licensing Agreements

The following conditions apply to the licensing of digital material:

- A written application must be made in advance to the Head of Reader Services Division to obtain permission to reproduce a record in print, on television or on the internet.
- Permission can be granted to distribute a copy of a record for a maximum period of 7 years.<sup>1</sup>

---

<sup>1</sup> Schedule 4, National Archives (Fees) Regulations, 2012

## **8. Copyright**

Copies of records are created for private research purposes. If digital images are required for exhibition, publication (hardcopy and online) or television, specific permission must be obtained and charges may be levied (See Copyright Guidelines).

- Applications should be made in writing to the Head of Reader Services Division;
- The full citation of the reference code of the archival documents to be digitised must be included;
- Details of the proposed exhibition, publication or media use must be given.

## **Procedures**

### **9. Preparatory work for digitisation**

Before a collection is digitised the following procedures must be followed:

- The collection must be formally accessioned and a National Archives reference code assigned;
- The Senior Archivist in charge of digitisation will carry out an assessment to decide if the documents will be scanned or photographed;
- The Senior Conservator will carry out a conservation assessment;
- The collection must be catalogued and each individual sheet and volume pages sub-numbered, including blank pages where appropriate;
- Conservation repair work will be carried out following listing and numbering.

### **10. Image capture**

The image capture work will be carried out as follows:

- Preservation image capture will be as TIFFs;
- For archival material up to and including A4 size material the resolution will be at 450DPI;
- For larger archival material the resolution will vary between 250 and 400 DPI TIFF depending on the physical format of the collection (e.g. map, volume, file etc.);
- A spreadsheet containing the item level references of the archival collection will be required to link the images to the descriptions;
- Digital images will be arranged in the same hierarchical order as the finding aid (fonds, series, item level);
- Digital images will be named with a similar reference number to the physical items;
- Access copies in JPEG or PDF format will be created from the TIFF images. For images of printed material optical character recognition will be applied to the PDF surrogates;
- Once image capture is complete the JPEG or PDF images will be linked to the spreadsheet using a concatenating formula.

## **11. Online access**

Online access to digital collections is currently provided through the various collection specific microsites on [www.nationalarchives.ie](http://www.nationalarchives.ie). In line with the National Archives strategic plan, it is intended that digitised collections will, at a future date, be stored in a digital repository which will facilitate both browsing and searching.

## **12. Storage and retrieval of digitised collections**

Access copies for each collection (JPEG/PDFs) will be stored locally on the National Archives network for six months after which they will be transferred to external tape storage.

Two copies of the TIFF images will be backed up onto external tape storage. The TIFF copies are stored in geographically distinct locations and will be periodically tested to ensure their continued viability.

## **Bibliography**

“Digitisation at the National Archives”, by The National Archives, Crown Copyright 2013, available at <http://nationalarchives.gov.uk/documents/information-management/digitisation-at-the-national-archives.pdf>

“Digital Imaging Standards Policy V.7”, by Tim Keefe, Digital Resources and Imaging Services, Trinity College Library, Dublin (June 2011), available at [http://dri.ie/sites/default/files/files/TCDImaging\\_Standards\\_v7.pdf](http://dri.ie/sites/default/files/files/TCDImaging_Standards_v7.pdf)

“Technical Guidelines for Digitizing Cultural Heritage Materials”, Federal Agencies Digitization Guidelines Initiative, Draft (September 2015), available at [http://www.digitizationguidelines.gov/guidelines/FADGI\\_Still\\_Image\\_Tech\\_Guidelines\\_2015-09-02\\_v4.pdf](http://www.digitizationguidelines.gov/guidelines/FADGI_Still_Image_Tech_Guidelines_2015-09-02_v4.pdf)

“Digitisation Policy”, by Public Record Office of Victoria, available at <http://prov.vic.gov.au/about-us/collaboration/digitisation/digitisation-policy>

# Appendix 1: Technical Standards

1. Introduction
2. Equipment
3. Image capture
4. File formats
5. Metadata
6. Processing
7. Workflow

## **1. Introduction**

This appendix details the technical standards to be followed during all image capture work undertaken by, and on behalf of, the National Archives.

## **2. Equipment**

National Archives approves the use of overhead scanners, scanners with a flat scanning bed, and cameras. Scanners with glass plates may be used.

All imaging equipment to be used by external contractors must be approved by the National Archives in advance.

## **3. Image Capture**

The objective of image capture is to produce an accurate reproduction of the original document, this takes priority over aesthetics. Every effort should be made to create the best digital surrogate possible.

- Digital surrogates must be legible and at least as readable as the original document;
- Images should be de-skewed and have a nominal skew of no more than one degree;
- Images should be single pages, unless otherwise agreed;
- All images should be assessed after scanning to check their quality and completeness. Documents to be rescanned if necessary;
- Digital images must demonstrate a similar organisational arrangement as original archival documents e.g. file structures and naming to mirror physical arrangement.

#### 4. File Formats

Two main types of files will be maintained, a master copy and a lower resolution surrogate of the image files. Surrogate image files will be generated directly from the master copy.

	Master Copy Image File	Surrogate Image File	Public Surrogate Image
<b>File Formats</b>	TIFF file format Compression – LZW (compression required to be lossless) Pixel Order – Interleaved Byte Order – IBM PC Image Pyramid – Unchecked <i>*This file will not be saved with any artefact generating, or “lossy” type compression</i>	JPEG (High Resolution) Image options – Quality 10 Format options – Baseline ‘Standard’	PDF (Low resolution)  <i>*PDF file conversion is the final step in the image processing workflow</i>
<b>Dimensional Characteristics</b>	Image scaled to 100% of original at time of capture  <i>*or at the highest device derived resolution from dedicated camera based imaging systems</i>	Image size, resolution and bit depth will be maintained from the master image	

#### 4.1 Compression

Lossless compression to be used for master digital images. Lossy compression is acceptable for lower resolution surrogate copies of master images.

#### 4.2 Resolution

The requirements as to Pixel per inch (PPI) will vary depending on the format of the material to be imaged.

- 300 PPI standard default for original records
- 600 PPI for photographs
- 300 PPI for photographic transparencies
- 300 PPI for microform, to be at the size of the original records

#### 4.3 Physical dimensions

- Digital images to have border to demonstrate that the entire page of a document has been captured.
- If it is not possible to capture the full page in one image, there should be enough of an overlap to reproduce the full page and show that no text has been lost.



#### 4.4 Bit depth (Signal Resolution)

	Master Archival Image File	Public Surrogate Image File
<b>Signal Resolution (Bit-Depth)</b>	<p>≥ 8 bit Greyscale</p> <p>≥ 24 bit Colour</p> <p>*highest bit depth capture at sensor level to be used, if image path processing up-samples the depth use 24 bit colour</p>	<p>8 bit Greyscale</p> <p>24 bit Colour</p>

#### 4.5 Colour

	Master Archival Image File	Public Surrogate Image File
<b>Colour Mode</b>	<p>Scanner output –</p> <p>LAB (if available as scanner native format)</p> <p>RGB (if LAB is not natively supported)</p> <p>*LAB output will be converted to RGB after processing and prior to master file save</p> <p>Greyscale</p>	<p>RGB</p> <p>Greyscale</p>

#### 4.6 Cropping

	Master Archival Image File	Public Surrogate Image File(s)
<b>Cropping</b>	<p>Digital image boundaries to extend for a minimum 1cm beyond boundary of physical item</p> <p>Cropping to be set to include colour and greyscale reference targets in Master archival image file</p>	<p>Duplicate the cropping of master archival image file</p> <p>In certain cases cropping that does not represent the entire physical item will be allowed</p>

#### 4.7 Backing

A solid backing material should be placed behind the document to be imaged to ensure a uniform field, and assist to define the boundaries of the original document.

#### 4.8 Rotation/Skew

All digital images should be straight.

	Master Archival Image File
Rotation/Skew	Rotation/skew $\leq 5^\circ$ Can be corrected through image editing software Rotation/skew $\geq 5^\circ$ Will require recapture of a new digital image

#### 4.9 Capture

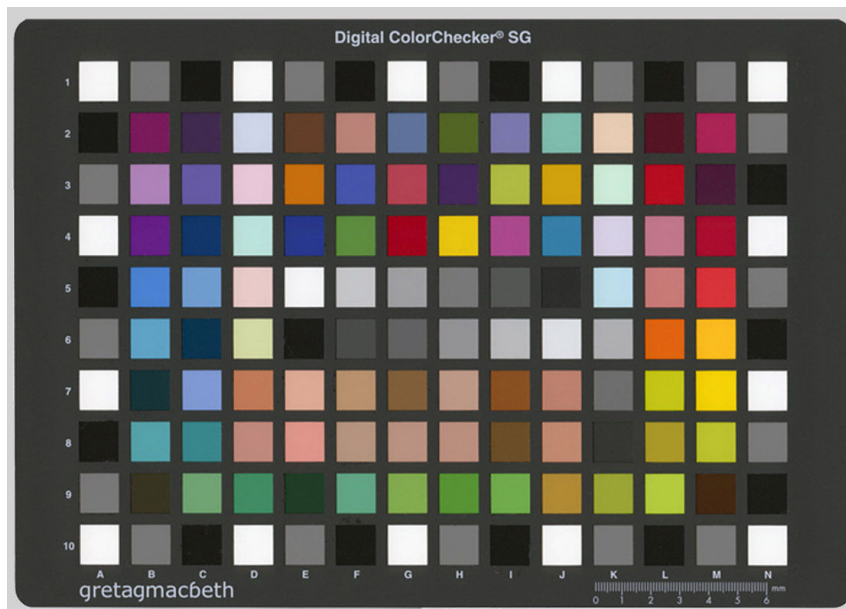
Digital images should be of the highest standard and quality to create an accurate record of an original document at the time of capture. Image capture work should not try and recreate an image of the document in its original form (as it looked when created).

#### 4.10 Calibration

Reference targets may be used to support photographic processes.

Greyttag Macbeth Digital ColorChecker SC

- This target to be used with the X-Rite iOne photo SG colour calibration package, for use in the characterisation of input devices.



#### 4.11 Imaging Equipment

- Imaging equipment (cameras/scanners) to be maintained to the manufacturers' recommended calibration and maintenance schedules and procedures.
- Computer monitors to be calibrated using i1Photo Pro 2.

	Frequency
Monitor Calibration	Twice monthly
Scanner Profiling	Twice yearly
Scanner Verification	Daily
Scanner Calibration	As per manufacturers recommendations

#### 4.12 Image quality

Quality checking will be carried out on a random sample of images from each imaging project. The criteria to be used for this checking will include:

- File naming structure
- Cropping
- Rotation/skew
- Sharpness
- Colour profile
- Accuracy of final image
- Accuracy of metadata

### 5. Metadata

National Archives requires that metadata be recorded for all digitisation work. While individual digitisation projects may have specific metadata requirements, in general, the following metadata elements shall be recorded for each digital image:

Reproduction reference	The unique file name and path of the image file described in this record.
Format	The file format, physical medium, or dimensions of the resource. National Archives shall use a controlled vocabulary of Internet Media Types (NIME), in line with international best practice.
Reproduction Type	Keyword to denote the type of the reproduction described here, e.g. TIFF. National Archives shall use a controlled vocabulary, the DCMI Type Vocabulary (DCMITYPE), in line with international best practice.
Copies	Use this field to enter the number of images, e.g. 1
Technique	Record the techniques used in the making of the reproduction, e.g. scanned image, digital photography.
Date	The date on which the reproduction was made.
Title	The title of the reproduction.

Creator	This field should be used to enter the name of the photographer or creator(s) of the reproduction. This may be the name of a person, group of people, or an institution.
Subject	The topic of the resource.
Description	Description may include but is not limited to: an abstract, table of contents or a free-text account of the resource.
Publisher	The entity responsible for making the resource available. This will generally be the National Archives, Ireland.
Contributor	An entity responsible for making contributions to the resource. Examples of a contributor include a person, organisation or service. Typically, the name of a contributor should be used to indicate an entity.
Source	A related source from which the described resource is derived. This will generally be the reference code of the original archival document.
Coverage	The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant. Typically, this will be Ireland.
Rights	Copyright and intellectual property rights which apply to the reproduction. Should also include a statement regarding the right to reproduce the image, i.e. 'Archives are subject to copyright and should not be copied or reproduced without the written permission of the Director of the National Archives'.
Notes	This field should be used to enter any additional notes about the reproduction.

## 6. Processing

Image processing work is carried out after the initial image capture is completed. Making a copy of the master TIFF file, and editing work can be carried out on the surrogate JPEG image. The image processing work to be performed is not to correct substandard digital image capture, but to address issues of legibility, e.g. ensure all text is legible. If imaging processing cannot amend any issues with the image, the decision can be made to re-take the image.

Adjustments that may be performed during processing:

- Physical adjustments – crop/rotation/skew adjustment
- Colour correction and tone adjustment
- Aim points
- Sharpening
- Artefact removal

## 6.1 Watermarking

A visible watermark may be applied to a surrogate digital image file for public use. There are two types of watermark:

- National Archives logo to be placed in centre of document, logo will be provided by National Archives.
- National Archives, Ireland accompanied by document reference, e.g. NAI/2040/1111/123.

## 6.2 Headers and footers

A header or footer may also be used, with the text agreed with the National Archives.

## 7. Workflow

An indicative workflow is as follows:

- **Records selection**
  - Assess original documents for digitisation
  - Scope and agree conservation requirements
  - Test imaging to agree digitisation by scanner or photography
  - Set timeframe for imaging work
  - Undertake focused pilot to validate workflow
  - Evaluate IT implications
- **Metadata Scoping**
  - Project naming
  - Agree metadata fields and data syntax.
- **Digitisation**
  - Original documents produced from archive repository and brought to Digital Imaging Unit
  - Imaging work carried out and throughput recorded
  - Original documents returned to archive repository
  - Image data transferred from scanner PC to processing PC
- **Processing**
  - Project name and file folders organised in same format as original archival documents
  - Surrogate generation
    - High resolution full size JPEG
    - Low resolution PDF
    - Apply watermark (if required)
    - Apply header/footer (if required)

- Imaging processing
  - Physical adjustments
  - Colour and tonal adjustments
  - Aim points
  - Sharpening
  - Artefact removal
  
- **Data backup**
  - Digital Imaging Unit Image Drives
    - TIFF archival master
    - High resolution JPEG
    - Low resolution PDF
  - National Archives Backup Servers
    - TIFF archival master
  
- **Metadata Processing**
  - Generate project headings and structure in Image Repository
  - Upload images to Image Repository
  - Input metadata
  - Quality assurance check on randomly selected images
  
- **Project close**
  - Clean data from scanning and processing computers
  - Notify relevant personnel in National Archives of completion of project
  - Undertake project review as necessary

## **Appendix 2: Conservation Requirements**

### **Preparing original documents for digital imaging**

A conservation assessment will be carried out on all documents to be imaged.

The conservation assessment will include a visual examination of all pages/folios to check that the text is legible and to ensure the documents are not too fragile for digital imaging. The assessment will also check for mould, pages stuck together and inserts or fastenings obscuring text.

### **Conservation work**

Once the conservation assessment is complete a programme of work will be established to prepare the documents. The following measures may be carried out when preparing documents for digital imaging.

- **Removal of surface dirt**
  - dirt is removed as it can obscure the legibility of text;
  - it may also cause further damage or harm to a document, e.g. if reverse of a page is very dirty it might soil adjacent pages.
  
- **Removal of creases and folds**
  - creases and folds obscure text and may give the impression of obscuring text on the surrogate image;
  - they inhibit safe handling;
  - they can be close to tears identified for repair.
  
- **Removal of metal fastenings**
  - metal fastenings are removed if they are made of, or contain, metal parts, e.g. pins, paper clips, staples or metal ended treasury tags to prevent future damage to adjacent documents;
  - they will be removed if they are causing major physical damage to the document, e.g. threads causing tearing of pages, 'sawing action' when turning pages;
  - they are obscuring text, e.g. if pages cannot be opened so that the entire text can be imaged;
  - they are loose and may get lost.
  
- **Repair of physical damage to the pages**
  - tears that are close to, or affect the text area will be repaired;
  - if the size or quantity of tears restrict safe handling of the whole document;
  - tears are along folds or weakened areas;
  - tears are longer than 2 centimetres.